

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号  
特開2000-244814  
(P2000-244814A)

(43)公開日 平成12年9月8日(2000.9.8)

(51)Int.Cl. <sup>7</sup>	識別記号	F I	テームコード <sup>8</sup> (参考)
H 0 4 N 5/265		H 0 4 N 5/265	5 B 0 5 7
G 0 6 T 1/00		G 0 6 F 15/66	4 7 0 J 5 C 0 2 3

審査請求 未請求 請求項の数10 O L (全 12 頁)

(21)出願番号 特願平11-45935

(22)出願日 平成11年2月24日(1999.2.24)

(71)出願人 000005108  
株式会社日立製作所  
東京都千代田区神田駿河台四丁目6番地  
(72)発明者 長坂 晃朗  
東京都国分寺市東恋ヶ窪一丁目280番地  
株式会社日立製作所中央研究所内  
(72)発明者 宮武 孝文  
東京都国分寺市東恋ヶ窪一丁目280番地  
株式会社日立製作所中央研究所内  
(74)代理人 100068504  
弁理士 小川 勝男

最終頁に続く

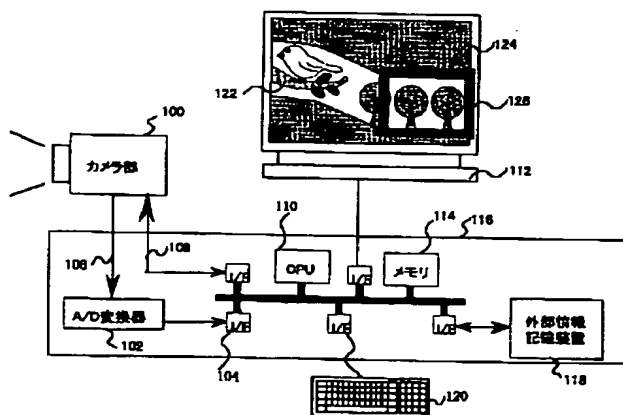
(54)【発明の名称】 画像合成装置、画像合成方法が記録された記録媒体

(57)【要約】

【課題】 本発明の目的は、動画からのパノラマ画像や高解像画像の合成において、次々と入力される大量の枚数のフレーム画像を遅延なくリアルタイムに処理し記録しながら、さらに、合成処理の進行状況や結果を示すインジケータ表示をも同時実行可能にすることにある。

【解決手段】 上記の目的を達成するため、撮影中の動画を表示する表示手段と、前記動画から順次フレーム画像を取り込み、取り込み時刻の異なる2枚のフレーム画像を比較して移動量を求める手段と、前記2枚のフレーム画像の差分画像及び前記移動量とを記録する手段と、記憶された移動量と取り込まれたフレーム画像とから合成された合成画像を象徴するインジケータ情報を作成する手段と、前記差分画像と前記移動量を用いてパノラマ画像を合成する手段とを設ける。

図1



## 【特許請求の範囲】

【請求項1】撮影中の動画を表示する表示手段と、前記動画から順次フレーム画像を取り込み、取り込み時刻の異なる2枚のフレーム画像を比較して移動量を求める手段と、前記2枚のフレーム画像の差分画像及び前記移動量とを記憶する手段と、記憶された移動量と取り込まれたフレーム画像とから合成された合成画像を象徴するインジケータ情報を作成する手段と、前記差分画像と前記移動量を用いてパノラマ画像を合成する手段とを有すること特徴とする画像合成装置。

【請求項2】請求項1に記載の画像合成装置において、前記インジケータ情報は、前記合成画像を縮小したものであって、前記取り込まれたフレーム画像を所定の倍率で縮小した縮小フレーム画像と前記移動量を該倍率に合わせて変換した移動量とから縮小画像を合成する手段を有し、前記表示手段は、前記縮小画像及び前記パノラマ画像を表示することを特徴とする画像合成装置。

【請求項3】請求項1に記載の画像合成装置において、前記インジケータ情報は、前記合成画像の全体形状を象徴するシンボル画像であって、前記合成画像の形状を示すシンボル画像を作成する手段を有し、前記表示手段は、前記シンボル画像及び、前記パノラマ画像を表示することを特徴とする画像合成装置。

【請求項4】請求項2に記載の画像合成装置において、前記縮小画像中における、最新に取り込まれたフレームに対応する領域を強調表示することを特徴とする画像合成装置。

【請求項5】請求項3に記載の画像合成装置において、前記シンボル画像中における、最新に取り込まれたフレームに対応する領域を強調表示することを特徴とする画像合成装置。

【請求項6】請求項1に記載の画像合成装置において、前記差分画像は、前記2枚のフレーム画像の共通画像領域も含み、前記パノラマ画像合成手段は、該共通画像領域を用いて前記2枚のフレーム画像間の境界を滑らかに合成することを特徴とする画像合成装置。

【請求項7】請求項1に記載の画像合成装置において、前記差分画像を拡大して合成画像を作成する場合は、拡大された該差分画像の画素を補間によって計算することを特徴とする画像合成装置。

【請求項8】撮影された動画を表示する表示手段と、前記動画から順次フレーム画像を取り込む手段と、取り込んだ時刻の異なる2枚のフレーム画像を比較して画像の移動量を求める手段と、前記移動量に基づき取り込んだフレーム画像を組み合わせて合成画像を得る手段と、前記合成画像もしくはその縮小画像を記憶するバッファ手段と、前記バッファに記憶された前記合成画像の一領域を表示するプレビュー表示手段と、前記合成画像もしくはその縮小画像が所定の領域を超えた場合に、前記バッファに記憶された該合成画像及び前記一領域とを縮小す

る手段とを有することを特徴とする画像合成装置。

【請求項9】撮影中の動画を表示するステップと、前記動画から順次フレーム画像を取り込み、取り込み時刻の異なる2枚のフレーム画像を比較して移動量を求めるステップと、前記2枚のフレーム画像の差分画像及び前記移動量とを記憶するステップと、記憶された移動量と取り込まれたフレーム画像とから合成された合成画像を象徴するインジケータ情報を作成するステップと、前記差分画像と前記移動量を用いてパノラマ画像を合成するステップとを有すること特徴とする画像合成方法が記録されたコンピュータ読みとり可能な記録媒体。

【請求項10】取り込んだ時刻の異なる2枚のフレーム画像を比較して画像の移動量を求めるステップと、該移動量に基づき取り込んだフレーム画像を組み合わせて合成画像を得るステップと、該合成画像もしくはその縮小画像をバッファに記憶するステップと、該バッファに記憶された該合成画像の一領域を表示するプレビュー表示するステップと、該合成画像もしくはその縮小画像が規定の領域範囲を超えた場合に、該バッファに記憶された該合成画像及び該一領域とを縮小して表示するステップとを有すること特徴とする画像合成方法が記録されたコンピュータ読みとり可能な記録媒体。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、広視野角のパノラマ画像を合成する合成装置及び合成方法が記録された記録媒体に係り、特に連続する動画像からリアルタイムにパノラマ画像を合成する画像合成装置、及び該合成方法画記録された記録媒体に関する。

## 【0002】

【従来の技術】異なる視野を撮影した複数の画像を合成して、1枚の広視野のパノラマ画像を作る技術が注目されている。最近では、特に、ビデオカメラをパンして視野を変えながら撮影した連続動画像から同様にパノラマ画像を作成する、Video Mosaicと呼ばれる技術に関心が集まっている。例えば、米国マサチューセッツ工科大学のTeodosio、L.、Bender、W.らによるSalient Video Stills: Content and Context Preserved、Proc. ACM Multimedia 93、pp. 39-46 (1993)などが知られている。この他にも様々な研究例が存在するが、いずれの手法も、連続して得られた複数枚の画像について、その共通領域を見つけ、ちょうど共通領域が重なるように、画像を貼り合わせることでパノラマ画像を得ている。カメラを動かすに伴って、撮影画像には新しい視野の画像が入ってくるので、それらが繋ぎ合わされて、最終的に広視野のパノラマ画像が得られる。

【0003】しかしながら、多くの研究例では、このような共通領域をいかに精度良く見つけるかや、レンズの歪みによる画像周辺部のずれをいかに補正するか、など、合成画像の精度を高めることが重視され、このよう

なパノラマ合成技術を実際に用いる操作者のユーザビリティの観点からの検討が十分とはいえない。一般に、対象となる映像の種類を特定できない自然映像について画像認識を行うには、様々な局面で、人間が介入し、適切な判断を与える必要がある。この場合、処理の高速応答性が、操作者の使い勝手を大きく左右する。用途によっては、画像として破綻なく接続されてさえいれば、むしろ入力映像をリアルタイム、即ち、映像の入力に対して遅延なく処理でき、その結果が直ちに確認できる、ユーザビリティに優れた手法のほうが望ましい場合も少なくない。

#### 【0004】

【発明が解決しようとする課題】パノラマ画像の作成においては、2種類の処理が行われる必要がある。一つは、連続する入力画像を認識して、共通領域の座標を決定する、カメラ動き量の推定処理である。もう一つは、得られた動き量から、実際に画像を繋ぎ合わせる合成処理である。ユーザビリティの観点からは、これらが同時に、しかもリアルタイムで行われることが最も望ましい。すなわち、カメラを振ると同時に、その振った量が計算されて自動的に画像が繋ぎ合わされ、撮り終わったときには合成画像がすでに出来ていて確認できることが理想である。これによって、気に入った画像が得られるまで何度でも気軽にやり直しができ、情報が欠落しているなどの本質的に合成が不可能な映像の撮り方をして、その場でミスがわかり、撮り方からの根本的な修正ができる。

【0005】カメラを振ると同時に、その振った量が計算されて自動的に画像が繋ぎ合わされ、撮り終わったときには合成画像がすでに出来ていて確認できるようにする場合、これら2種類の推定処理、合成処理のいずれもリアルタイムに処理を行う必要がある。カメラ動き量の推定処理に関しては、例えば、発明者らによる特開平11-004398「デジタルワイドカメラ」で示した高速処理方式などがある。しかしながら、合成処理については、リアルタイムに同時に実行できるレベルにまで十分に検討されているとはいえない。単に、動き量を撮影と同時にリアルタイムで計算できるだけでは、撮影後に、動き量の計算分は短くなるものの、画像どうしの貼り合わせ処理の時間を待たねば、合成結果を確認することはできず、ユーザビリティの面で不十分である。

【0006】本発明の目的は、連続する動画像からパノラマ画像を合成する方法において、操作者のユーザビリティを向上させるため、合成結果をリアルタイムで確認できる、合成処理の高速応答性を実現することにある。

#### 【0007】

【課題を解決するための手段】上記の目的を達成するため、撮影中の動画を表示する表示手段と、前記動画から順次フレーム画像を取り込み、取り込み時刻の異なる2枚のフレーム画像を比較して移動量を求める手段と、前

記2枚のフレーム画像の差分画像及び前記移動量とを記録する手段と、記憶された移動量と取り込まれたフレーム画像とから合成された合成画像を象徴するインジケータ情報を作成する手段と、前記差分画像と前記移動量を用いてパノラマ画像を合成する手段とを設ける。

#### 【0008】

【発明の実施の形態】以下、本発明の一実施例を詳細に説明する。図1は、本発明を実現するためのシステム構成の概略ブロック図の一例である。100は一般的なビデオカメラ撮像部であり、撮影した映像を映像信号線106を介して演算処理装置116に送信する。カメラ部は、制御信号線108によって演算処理装置116と接続され、116は必要に応じてビデオカメラのズーム倍率を変更したり、ホワイトバランスを調整したり、といったカメラパラメータの設定ができるようになっている。このような設定が不要な場合は、省略可能である。演算処理装置116は、現在汎用的に用いられているデジタルコンピュータ、特にパソコンのシステム構成と同じであり、そのディスプレイ出力は映像信号線を通じてディスプレイ112の画面上に描画され、入力デバイス120でユーザから受けた指示は、演算処理装置116に通知される。カメラ部100から出力される映像信号は、逐次、102のA/D変換器によってデジタル画像データに変換され、インタフェース104を介してメモリ114に入り、メモリ114の別の番地に格納されたプログラムに従って、CPU110によって処理される。

【0009】ここでは、ビデオカメラからの信号がNTSC等のアナログビデオ形式の場合の例で説明しているが、デジタル出力のビデオカメラを利用する場合には、102のA/D変換器は不要である。この場合、代わりに画像形式の変換器が必要になる場合もある。また、処理の必要に応じて、各種情報を118の外部情報記憶装置に蓄積することができる。

【0010】メモリ114には、以下に説明する処理によって作成される各種のデータが格納され、CPU110が処理を行う際に、必要に応じて参照される。以下の説明では、すべてCPUが実行するソフトウェアとして説明するが、その一部または全部をハードウェア論理回路で置き換えて実行させても構わない。また、カメラ部100、演算処理装置116、ディスプレイ112を一体化したインテリジェントなビデオカメラ形態であっても構わない。

【0011】以下では、上記ハードウェア上で実行される、先述の課題を解決するソフトウェアフローについて詳細に説明する。まず最初に、上記ソフトウェアを用いた、パノラマ画像の作成方法について説明する。ユーザは、最初にリアルタイムでパノラマ合成が行われる状態（以下では、パノラマ撮影モードと呼ぶ）に移行することを、入力デバイス120によって演算処理装置116に伝える。入力デバイスは、コンピュータシステムを利用する場合には、キーボードやマウスなどにあたり、カ

メラ一体型形態をとる場合には、「パノラマ撮影開始ボタン」などのスイッチが相当する。演算処理装置116は、パノラマ撮影モードに入ると、ビデオカメラから出力される映像信号を毎フレーム入力し、入力されたフレーム画像のカメラ動き量を求め、パノラマ合成に必要なパラメータとデータを収集する。このとき、ディスプレイ112には、その全体、もしくは一部の領域が、合成状況を確認するためのプレビュー画面として設定され、カメラの動きに合わせて繋ぎ合わされていく合成画像の様子をリアルタイムで更新表示する。

【0012】図1のディスプレイ112内に、そのイメージを示す。本発明では、プレビュー表示用の合成画像122（以下、プレビュー画像と呼ぶ）は、プレビュー画面内で、その全体がおさまる最大の大きさになるように常にスケーリングされて表示されることを特徴とする。これによって、全体の合成状態を、瞬時に把握可能にする。このとき、プレビュー画面が矩形の場合には、合成画像の形状によって、画像の存在しない空白部分124が存在するが、これは適当な背景色で埋めておく。合成画像中における最新の入力フレームの位置は、枠で囲んだり、色を変えたりして強調表示することで、カメラを動かすべき方向を迷わせない羅針盤の役割も持たせる（126）。

【0013】ユーザが、再び入力デバイス120によって、パノラマ撮影モードの終了を指示すると、新規のフレーム画像入力を中止し、待機状態になる。ディスプレイ112には、最後のフレーム入力によって更新されたプレビュー画像が保持され、ユーザはその画像を見て所望の結果が得られているか判断し、良ければ、プレビュー用ではない実サイズの最終的な出力としての合成画像（以下、最終出力合成画像と呼ぶ）を作成し、外部情報記録装置118などに保存する。

【0014】図2に、上記機能を実現するフローチャートの一例を示す。このアルゴリズムは、上記機能に必須の高速処理を可能にするため、次の3点を基本コンセプトとしている。

【0015】(1) 最終出力合成画像とは別に、プレビュー専用縮小サイズの合成画像を作成する。

【0016】(2) 最終合成画像用に記憶する画像は、最低限必要な領域のみとする。

【0017】(3) 任意形状・任意サイズのプレビュー画像の作成と、メモリ効率とを両立させる。

【0018】(1)は合成処理にかかる時間を減らすための工夫である。一般に、作成された画像が所望のものかどうかを判断する目的では、必ずしも大きなサイズの画像である必要はない。そこで、撮影時には、最終出力合成画像は作成せず、プレビュー用として別個に縮小サイズの合成画像を作成する。これによって、撮影とほぼ同時にプレビュー画像が完成するようにする。このプレビュー画像は、得られたカメラ動き推定量に基づいて合

成が行われるので、実際の合成結果に極めて近く、歪みやミスなどもそのまま反映される。ユーザは、このプレビューによって所望の画像が得られたか否かを判断し、良しと確認されたら、改めて実際の大きさの最終出力合成画像を作成する。失敗と判断した場合には、直ちにやり直すことができ、実サイズの合成の時間を省けるので、効率的な作業ができる。

【0019】(2)は、最終出力合成画像のための元データの記録時間を削減するための工夫である。(1)を採用して最終出力合成画像の作成を後で行うようにした場合、その元データであるフレーム画像を一時的に記録しておく必要がある。次々と入力される新しいフレーム画像を次々に記録していくことは、データ転送時間が長かったり、必要なメモリ容量が肥大化したりと、高速処理の阻害要因となる。

【0020】ここで、図3に示すように、パノラマ合成画像は、連続するフレーム画像の1枚1枚を重ね合わせて得られる。図3は、一連のパン映像からパノラマ画像を作成する例であり、上側が連続するフレーム画像からなる映像を模式的に表し、下側がそれから合成されるパノラマ画像を示している。上下を結び付ける複数の線は、フレーム画像どうしの対応関係を表している。下側のパノラマ画像では、連続するフレーム間で常に画像的に一致する領域（以下、共通領域と呼ぶ）が存在し、データとして冗長性が極めて高くなっている。そこで、この連続フレーム間の共通領域を省いた部分の領域だけを組み合わせることで、十分にパノラマ画像が再現できる。従って、1フレームにつき、この共通領域を省いた領域だけを記録するようにすれば、1フレームあたりの記録すべきデータ量は大幅に削減され、データ転送時間は短縮され、必要なメモリ容量も少なくなる。最後のフレームの重ねあわせだけでは画像全体が必要になるが、唯一その1回のみであり、また、最後であるから、データ転送に時間が多少多くかかっても、次のフレームが存在しないため、次フレームのための処理に間に合わないという問題も発生しない。

【0021】最後の(3)についてであるが、プレビュー画像は、撮影終了と同時に得られるだけでなく、さらに撮影中からカメラの動きに即応してリアルタイムに、即ちカメラの動きに同期して更新されるほうが、ユーザにとって確認が容易であり、使い勝手がよい。しかし、合成されるパノラマ画像は、カメラを動かすにしたがって単調に大きくなり、また、カメラの動かし方によって、その形状や大きさは様々になる。プレビュー画像を表示するためには、プレビュー用に合成されるパノラマ画像が常にメモリ領域上で作成されていなければならないが、可能性のあるすべての形状や大きさのプレビュー画像をメモリ領域上に作成するためには、メモリ領域を予め極めて大きくとっておかなければならない。

【0022】(1)の指針により、縮小フレーム画像を用

いるとしても、必要なメモリ領域の容量は巨大である。しかも、実際には利用されないメモリ部分も多く、メモリの使用効率が非常に低くなってしまふ。もちろん、最初にカメラを動かす方向を決定して、それに必要な分だけメモリ領域を確保することはできる。しかし、この方法では、ユーザが予めカメラを動かす方向を指示しなくてはならない煩雑さが生じ、ましてユーザが途中で気が変わっても全く対応できない。また、最終的にどれだけの大きさの画像になるか、ユーザが予測できない場合も多い。無制限のサイズ・無制限の形状の合成画像に対応することが望まれる。

【0023】ここで、モニタにプレビュー画像を表示するためのプレビュー画面の大きさは固定であり、プレビュー画像の大きさが固定の大きさのプレビュー画面より大きくなった場合、その画面サイズに合うように縮小スケールされて表示されるようにする。逆にいえば、プレビュー画面の解像度を超える画素数は、プレビュー画像を表示する際には無駄になる。

【0024】そこで、プレビュー画像の大きさが、プレビュー画面の大きさの $n$ 倍（縦もしくは横の長さとしての比率。面積的には $n \times n$ 倍）を超えた時点で、プレビュー画像の大きさを強制的に $1/m$ 倍（同じく面積的には $1/(m \times m)$ 倍）に縮小する。そしてまた、縮小したプレビュー画像の大きさが再びプレビュー画面の $n$ 倍を超えたときには、プレビュー画像の $1/m$ 倍縮小を行うというように繰り返す。これにより、プレビュー画像用のメモリ領域サイズを、プレビュー画面のサイズの $n$ 倍以下に制限でき、かつ、プレビュー画面に表示されるプレビュー画像は、それ以上のメモリ領域を持っている場合と比べて解像度の劣化がないという効果が得られる。

【0025】以上のことを踏まえて、図2のフローチャートの説明に戻る。まず各種の初期化を処理200で行う。例えば、プレビュー画像作成用のワーキングバッファとして、プレビュー画面の大きさを縦横それぞれ $n$ 倍にした大きさのメモリ領域（以下、プレビューバッファと呼ぶ）を確保する。プレビュー画像は、このプレビューバッファの中で合成される。プレビューバッファには、プレビュー表示領域と呼ばれる領域が設定され、この領域が、ちょうどプレビュー画面の大きさいっぱいになるように常にスケール、即ち、拡大・縮小されて、プレビュー画面の表示となる。

【0026】最初は、バッファ中央部の、プレビュー画面と同じ大きさの領域が、プレビュー表示領域として設定され、プレビューバッファ上でプレビュー画像が大きくなるにしたがって、常にプレビュー画像全体が含まれるように、順次プレビュー表示領域を拡大する。次いで、プレビュー画像作成用の変数である、プレビュー縮小率 $r$ の初期値を計算する（202）。 $r$ は、カメラから入力されるフレーム画像の大きさと、プレビュー画面の大きさとの比率であり、カメラの動きがない初期状態

において、入力フレーム画像が、プレビュー画面いっばいに表示されるように調整するパラメータである。例えば、入力フレーム画像の大きさとプレビュー画面の大きさを同じにした場合には、 $r=1.0$ となる。

【0027】続く処理204では、ビデオカメラで撮影された映像の最新の1コマのフレーム画像を入力する。そして、次々と入力されるフレーム画像間で比較を行いながら画像の移動量検出処理を行う。この移動量検出には、オプティカルフロー等の様々な方法が考えられているが、ここでは高速に計算可能な方式として、例えば、特願平9-153303の「デジタルワイドカメラ」に記載された方式を利用する。この方式は、画像の投影分布から移動量を求めるために非常に高速処理を可能とする。

【0028】処理206では、このフレーム画像から水平投影分布 $P_x$ 、垂直投影分布 $P_y$ を作成する。水平投影分布とは、画像を構成する画素の各行について、その色や輝度の平均値をとったもので、行数分だけの平均値の一次元系列である。垂直投影分布は、画像の各列について、同様に色や輝度の平均値をとったもので、列数分だけの平均値の一次元系列である。例えば、垂直投影分布の値の並びは、画像が左右に移動すると、それに対応して順方向もしくは逆方向にシフトする性質がある。同様に垂直投影分布は、画像の上下動に対応して順逆方向にシフトする。したがって、あるフレーム画像から作成した水平／垂直投影分布と、次に入力されたフレーム画像から作成した水平／垂直投影分布とを比較し、その平均値の系列がそれぞれどれだけシフトしたかを求めれば、2枚の画像の間の水平／垂直方向の移動量が求まることになる。画像変化なので、ノイズや若干の変形等により投影分布の値の並びは完全に一致するわけではないが、比較する2つの投影分布の片方を順次 $\delta$ 画素分だけシフトして比較し（ $\delta$ は $-R$ から $+R$ の範囲をとるすべての整数。 $R$ は予め設定した探索範囲を示す値）、最も相違度が低くなったときの $\delta$ を移動量として得る。相違度は、対応する行または列の画素の色や輝度の平均値それぞれの差の絶対の総和などを利用する（210～212）。但し、初めての画像入力の場合には、過去に比較するフレーム画像が存在しないので、処理204に戻る（208）。

【0029】処理214では、得られた水平／垂直移動量 $\delta$ から、直前のフレーム画像と現在のフレーム画像との共通領域を求め、直前のフレームにおける、その共通領域以外の部分の領域についてのみ記録する。これによって、全ての領域について記録する場合に比べ記憶容量が少なくて済む。

【0030】また、処理216では、得られた水平／垂直移動量 $\delta$ の値に、 $r$ を乗じたプレビュー合成用の移動量を求め、直前にプレビューバッファ上に描画した縮小フレーム画像に対して、上記で得られた $\delta \times r$ だけずらした位置に $r$ 倍に縮小した画像を重ねて描画する。この

ようにして拡大されたプレビュー画像全体を含むように更新したプレビュー表示領域を、プレビュー画面に表示する(218)。

【0031】このとき、もしプレビュー画像の全体が、プレビューバッファの端いっぱいには迫っているときには(220)、プレビューバッファ全体を $1/m$ 倍に縮小する(222)。このとき、プレビュー表示領域も同様に $1/m$ 倍に縮小される。そして、 $r$ も同様に $1/m$ を乗じておく。

【0032】図4に、このプレビューバッファの縮小処理の模式的な流れを示す。400は、プレビューバッファ全体を示し、402は、プレビュー表示領域である。404は、プレビュー画像を示す。プレビュー表示領域は、常にプレビュー画像全体に外接するように位置や大きさが変化する。Aは、パノラマ撮影開始直後の様子を表しており、入力されたフレーム画像が、中央部に描画されている。カメラを動かすにしたがって、プレビュー画像は伸びていき、プレビュー表示領域も追従して大きくなる(B)。そして、Cに示すように、プレビュー画像がプレビューバッファの端まで到達すると、バッファからはみ出すので、ここで $1/m$ 倍にプレビューバッファ全体を縮小する(D)。このとき、プレビュー表示領域も同じ倍率で縮小する。そして、さらにプレビュー画像が伸びれば、それに合わせて画像の描画を続ける。

【0033】ここで、 $n=3$ 、 $m=2$ とすると、はじめに、プレビューバッファの中央と表示領域の中央を合わせておき、その後、撮影画像に応じて伸び、端に到達したときのプレビュー表示領域のサイズは、プレビュー画面のオリジナルサイズの2倍(縦横の長さ比が2倍)よりも常に大きくなる。したがって、これをプレビューバッファ全体ごと $1/2$ 倍(縦横の長さ比を $1/2$ 倍)にしても、プレビュー表示領域のサイズは、プレビュー画面のオリジナルサイズを下回ることはない。なお、 $n=3$ 、 $m=2$ のとき、プレビュー表示領域がちょうどバッファの端に到達したとき、プレビュー表示領域は、プレビュー画面のオリジナルサイズの2倍となる。そこで、プレビューバッファを $1/2$ 倍すれば、プレビュー表示領域のサイズは、プレビュー画面のオリジナルサイズと同じになる。

【0034】これによって、巨大なメモリ領域を確保して、自由にプレビュー画像を描画し、その上で、プレビュー画面のサイズに間引いて表示する場合と、ほとんど変わらない結果が、プレビュー画面の $n$ 倍のメモリ容量を用意するだけで得られることになる。また、 $m=2$ の場合、単純に1画素おきに間引けばよいので、縮小も極めて高速に行える。

【0035】ここで、例えば、入力フレーム画像の10倍の大きさのパノラマ画像を作成することを考える。上記の方法によらずに、あらかじめメモリ領域を確保する方法では、この面積のパノラマ画像のプレビューを行う

ためには、あらかじめ、入力フレーム画像の約400倍のメモリ領域を必要とする。これは、最初からどの方向にカメラが動くか予測できないために、最初の入力フレーム画像を中心にして、上下左右の各方向にそれぞれ10倍の長さを持つ矩形領域を確保する必要があるためである。ところが、本発明の方法によれば、 $n=3$ の場合で、高々9倍の大きさで良い。本発明は、この9倍の大きさで無制限のサイズを保証するので、従来方法とのメモリ領域の差は、合成するパノラマ画像の大きさが20倍、30倍になるに伴い、指数関数的に大きくなる。

【0036】さらに、上記の手法は、ズーム映像を用いた高解像静止画合成処理の場合に、さらに効果がある。図5に、ズームによる高解像静止画合成の原理を示す。この基本的な考え方は、すでに挙げた文献Salient Video Stills: Content and Context Preservedに示されている。最もズームインされたときのフレーム画像を基準サイズとして、他のフレーム画像をそのズーム率に合わせて重ね合わせていくことで、大きな高解像度画像を得ることができる。この高解像度画像は、周辺部は画素が単純に大きくなるだけであるが、中心部に近づくにしたがって少しずつ段階的に高精細になる。通常のスチルカメラで撮影した画像も、中心に焦点を合わせたときには周辺部はボケており、重ね合わせ合成された高解像度画像の特性に近い。すなわち、安価な低解像度の撮像素子を用いて、高価な高解像度の撮像素子を用いて写したのと同等の画像が得られることになる。

【0037】しかし、このような従来の手法では、ズームアウト時の画像を拡大してから重ね合わせ処理を行わなければならない、拡大率が大きくなるにつれて拡大処理時間も長くなり、パノラマ合成のときよりもさらにリアルタイム処理が困難になる。それでも、徐々に単調にズームアウトするような映像の場合には、最初に巨大なメモリ領域が用意できれば、最初のフレームを原寸大で描画し、後はフレーム画像を拡大しながら重ねていくだけでよい。しかし、逆に、ズームインするような映像の場合は、高解像画像合成の目的である、解像度を高めながらの合成には、最後に入力されたフレーム画像が常に原寸大になるように、合成画像のほうを拡大する必要がある、巨大なメモリ領域を毎フレームごと拡大するのは非常に時間がかかる。このことは、予めどのようなカメラの動きがあるか予測できないリアルタイム処理ならではの課題である。

【0038】一方、本発明の考え方を適用すれば、少ないメモリ領域でリアルタイムにプレビューが確認できる。ズーム率を高速に求める手法については、例えば、すでに発明者らによる特願平10-53100「超解像カメラ」において示している手法が利用できる。ここで確認しておく、高解像画像の合成とは、連続するフレーム画像間の共通領域が同じ縮尺になるようにサイズを調整しながら、入力されるフレーム画像を重ね合わせていくこと

である。したがって、パノラマ合成のときと同様に、最終的な合成画像を得るには、共通領域は1回だけ記録すればよく、ズームインの映像の場合には、直前のフレームにおける、共通領域を除いた口の字型の外枠部分だけでよい。また、プレビュー画像の作成にあたっては、ズーム率に合わせて縮小したフレーム画像を、ただプレビューバッファに次々と描画していただくだけでよい。

【0039】ズーム率は、ここでは、最初に入力されたフレーム画像を基準にして、共通領域が一致するようにサイズ合わせをしたときの比率として定義する。ズームのないときは1.0、ズームインの場合は1より小さく、ズームアウトの場合、1より大きくなる。一方、ズームアウトの映像の場合には、最新のフレームのほうが広い範囲を写しているため、今度は最新のフレームにおける共通領域を除いた口の字型の外枠部分を記録する。プレビュー画像の作成にあたっては、この外枠部分をズーム率に合わせて縮小して描画する。ズームアウトの場合、プレビュー画像のサイズが、ズーム率が大きくなると、プレビューバッファのサイズを超えるので、その場合には、パノラマ合成のときと同様に、プレビューバッファ全体を1/nに縮小する。このように、ズームの場合にも、プレビュー画像作成のために必要な拡大は、口の字型の外枠部分だけをn倍以下に拡大する処理だけであり、明らかに高速に処理できる。以上によって、撮影と同時に、高解像画像合成の結果もリアルタイムで確認できる。

【0040】本実施例において最終的な合成画像のために記録される元データの一例を図6と図7に示す。データは、パノラマファイル構造体600とパノラマクリップ構造体700とに大別される。前者のパノラマファイル構造体は、元データ全体をまとめるヘッダーであり、後者のパノラマクリップ構造体は、各フレームごとに記録が決定される画像データである。パノラマファイル構造体には、後続データが画像合成用の元データであることを示す識別子として、ファイルID602と、合成画像の実サイズを幅(604)、高さ(606)の順に格納する。合成画像の最終的なサイズは、カメラ動き量から決定され、特殊な形状をとっていても、それに外接する最小の矩形として表現できるので、幅と高さのみを格納する。また、元データの色数608(例えば、256色なのか、24ビットフルカラーなのか等)や画像フォーマット610(例えば、RGB形式なのか、YUV形式なのか等)も格納する。

【0041】また、合成画像に外接する矩形領域の始点座標をパノラマ画像x、yオフセット(612、614)として格納し、この座標を基準にして、後述するパノラマクリップの座標を計算する。このヘッダー情報に続いて、パノラマクリップ構造体700がデータとして並ぶ。パノラマクリップ構造体は、フレーム1枚ごとのデータとなっており、合成画像における、そのフレーム

画像の位置とサイズを格納する。まず702は総データサイズであり、画像データの量を格納する。画像始点x座標702は、前述のパノラマ画像xオフセット612で示された座標を合成画像の始点座標としたときの、このパノラマクリップ構造体と対応づけられたフレーム画像の始点のx座標値を表す。同様に、画像始点y座標704は、フレーム画像の始点のy座標値を表す。画像幅706と画像高708は、フレーム画像の入力されたままの実サイズを表す。これに、ズーム率710を乗じて、合成画像上でのフレーム画像のサイズを計算する。714~720は、このフレーム画像における、前述の共通領域の範囲を指定する。共通領域始点x座標714と共通領域始点y座標716は、フレーム内における共通領域の始点座標を表し、共通領域幅718と共通領域高720で、その大きさを表す。画素データ722には、左上から右下方向にスキャンしながら、すべての画素データが順番に並ぶ。714~720に適切な値が入っている場合には、共通領域に関する画素データは単純に無視される。

【0042】次に、上記で記憶した元データから最終的な合成画像を作成する方法について説明する。まずパノラマファイル構造体600を参照して合成画像の幅と高さを調べ、その大きさの矩形画像領域をメモリ領域に確保する。そして、後続するパノラマクリップ構造体700を参照して、描画すべき位置と大きさを調べて、確保したメモリ領域内における相対位置を求め、画像データ722を書き込む。但し、ズーム率が1倍を下回るクリップが1つでも存在する場合には、最初に、最も小さなズーム率が1になるように、すなわち、最もズームインされたフレーム画像が等倍で描画されるように、すべてのパノラマクリップ構造体のズーム率を比例変換する。この場合、合成画像のサイズも比例して大きくなる。これによって、フレーム画像が縮小されることがなくなり、解像度が高められる方向にのみ合成が行われる。

【0043】以上で述べた画像合成をより美しく見せるために、スムージング処理を行うことができる。一般に、画像を貼り合わせる場合には、動き量が正しく求まらなかったり、レンズの収差による入力画像の時点での歪みがあったり、あるいは撮影時点での明るさが変動するなどして、画像の繋ぎ目に微妙な境界線が発生することがある。人間の知覚では、この境界線は不快な印象を与えるため、補正を行う必要がある。このような補正を行う最も簡単な方法は、貼り合わせる画像どうしの間に糊代を持たせ、境界線を挟んで、一方の画像の色から他方の画像の色へと段階的に推移するように貼り合わせることである。その概念を非常に単純化して示したイメージを図8に示す。800では、2枚の画像が単純に貼り合わされている様子を示している。ここでは、説明を簡単にするため、2枚の画像はそれぞれ特定の輝度で均一と仮定している。このときの水平方向の輝度の変化を見

てみると、画像の境目では、グラフ802に示すように、瞬時に変化し、これが境界線808を形成する。一方、804および806に示すように、画像の境界に糊代を持たせて、一方の画像の輝度から他方の画像の輝度まで段階的な変化となるように貼り合わせれば、境界線は目立たなくなる。

【0044】本発明の合成方法で、上記のスムージング処理を実現するには、元データを記録するときに、糊代の分だけ広めの領域を記録するようにすればよい。これによって必要最低限のデータ量の増加で、スムージングが可能になる。そして、後の合成処理の際に、糊代の分について、スムージング処理を行うことで、滑らかに接続された合成画像が得られる。糊代とする領域は、基本的に、連続する2枚の画像の間の共通領域を狭める形で選ぶ。図9に、2枚のフレーム画像900、902について、共通領域が重なるように配置した場合の例を示す。902が時間的に新しいフレーム画像であるとする、図左上のL字型ハッチング領域904が元画像データとして記録される。糊代は、2つの画像の境界線を段階的に埋めればよいので、境界線に沿って、共通領域を狭める方向に幅を膨らませた領域が最低限必要な領域となる。図9の例では、糊代となる領域は、904の内側のL字型のハッチング領域906となる。図10に、典型的なカメラ動き量と、そのときの糊代領域との組み合わせを一覧にして示す。図中、 $\Delta x$ 、 $\Delta y$ は、それぞれ水平／垂直方向の移動量を示す。図11に、さらに、ズームがある場合の同様の糊代領域のパターンを示す。図では、ズームインの場合を示している。このように、ズームがあっても、共通領域の大きさが同じになるように大きさを合わせたときの画像間の境界線から、共通領域方向に拡張して糊代領域を確定する。ズームアウトの場合も、フレーム画像900と902の関係が逆になるだけで基本的に同じである。

【0045】なお、高精細画像を作成する場合、ズーム率に合わせて画像データを貼り合わせる際に、単純な画素の拡大ではなく、バイリニア補間等の画素補間技術を利用して拡大することで、さらに見栄えのよい画像が得られることはいうまでもない。そして、上記のスムージング処理を、画素補間で拡大した画像データを対象に行うことで、さらに美しい画像が得られる。

【0046】また、上記では、リアルタイムでプレビュー画像を作成し表示する手法について述べたが、プレビュー画像の代わりに、合成中の様子を示すシンボルで簡略表示するようにして、ユーザビリティと、一層の高速化との両立を図ることもできる。例えば、プレビュー画像そのものではなく、そのシルエット、すなわち合成画像がどのような形状でできているかを線画で示すことに

より、フレーム画像を縮小して貼り合わせることも必要がなくなると同時に、線画のみなので高速に描画が可能になる。図12に、これに基づくプレビュー画面の例を示す。プレビュー画面には、入力映像がそのまま表示され、それと合わせて、線画1200もインジケータ表示される。また、現在入力中のフレームが合成画像中のどこに当たるのかを示すため、1202に示すように、線画の中で現在のフレーム画像が占める領域を強調表示してもよい。これによって、合成状況をリアルタイムで確認するという目的が、より低コストの演算性能で実現される。

#### 【0047】

【発明の効果】本発明によれば、パノラマ画像や高解像画像の作成において、撮影と同時に、処理の進行状況や結果を確認でき、気に入った画像が得られるまで、何度でも気軽に撮り直しが可能な、優れたユーザビリティを提供できる。

#### 【図面の簡単な説明】

【図1】本発明を実現するためのシステム構成図の一例である。

【図2】本発明を実現するための処理のフローチャートの一例である。

【図3】パンする動画からパノラマ画像を合成するイメージ図である。

【図4】本発明を実行した場合のプレビューバッファの内容の一例である。

【図5】ズームする動画から高解像画像を合成するイメージ図である。

【図6】本発明の処理で用いられるデータ構造体の一例である。

【図7】本発明の処理で用いられるデータ構造体の一例である。

【図8】画像貼り合わせ方法を示すイメージ図である。

【図9】本発明の処理で定義した各画像領域を説明する図である。

【図10】本発明の処理で用いられる糊代領域のパターンの一例である。

【図11】本発明の処理で用いられる糊代領域のパターンの一例である。

【図12】本発明の処理で用いられるインジケータ表示の一例である。

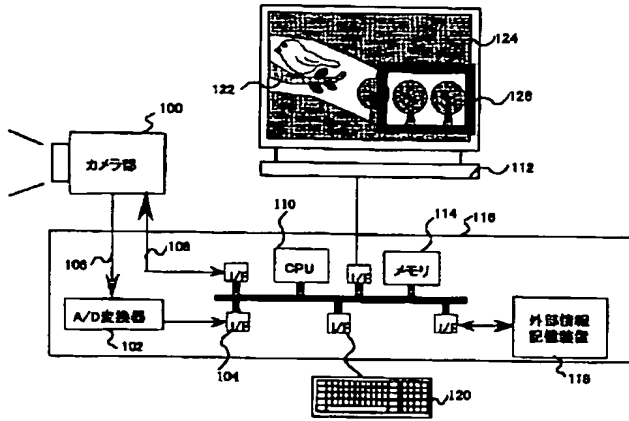
#### 【符号の説明】

100…カメラ部、102…A/D変換器、104…インタフェース、106…映像信号線、108…制御信号線、110…CPU、112…ディスプレイ、114…メモリ、116…演算処理装置、118…外部情報記憶装置、120…入力デバイス。



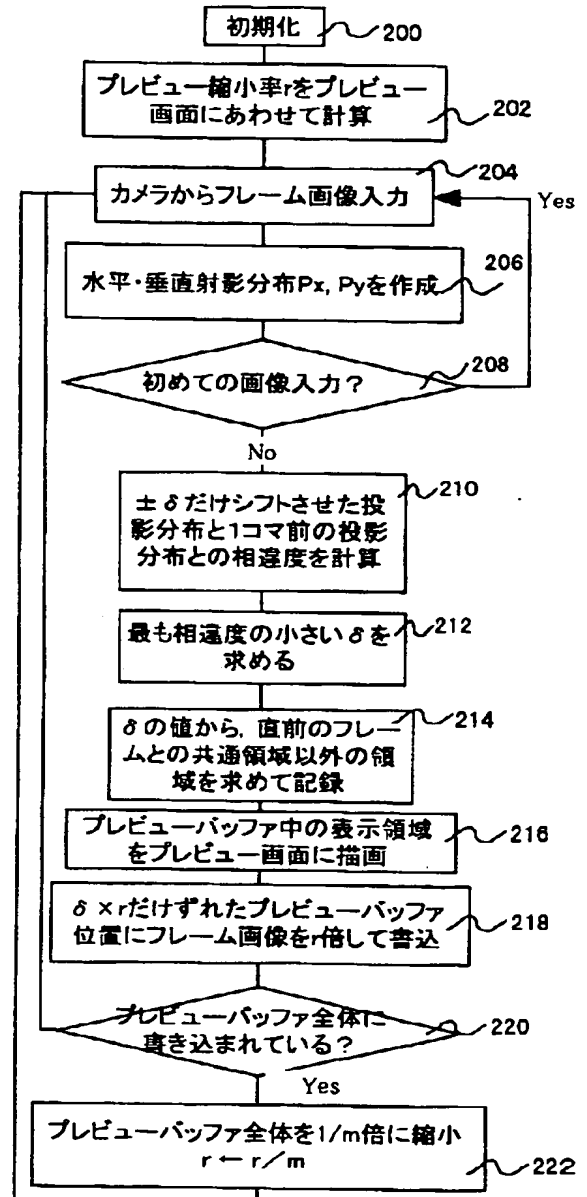
【図1】

図1



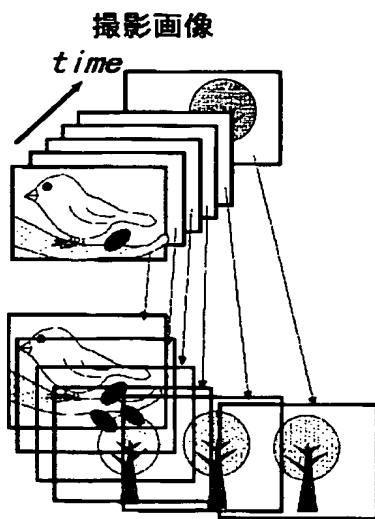
【図2】

図2



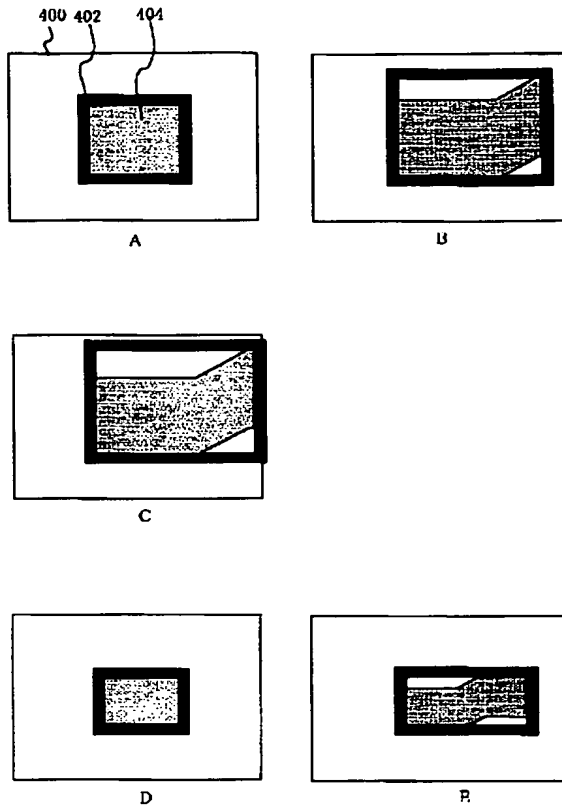
【図3】

図3



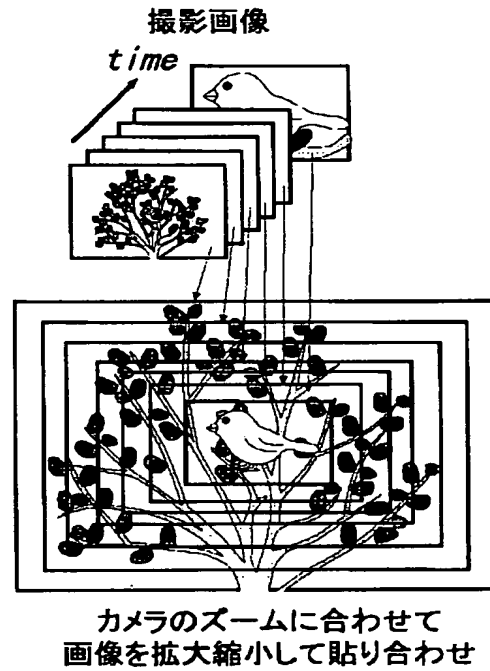
【図4】

図4



【図5】

図5

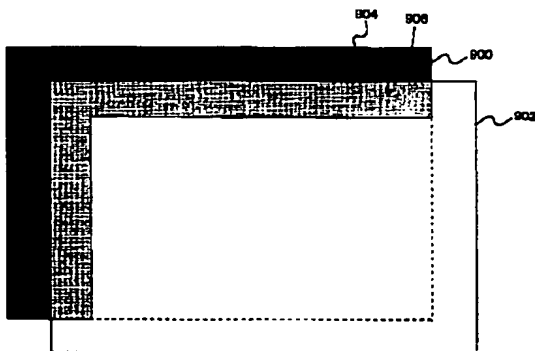


【図6】

図6

【図9】

図9

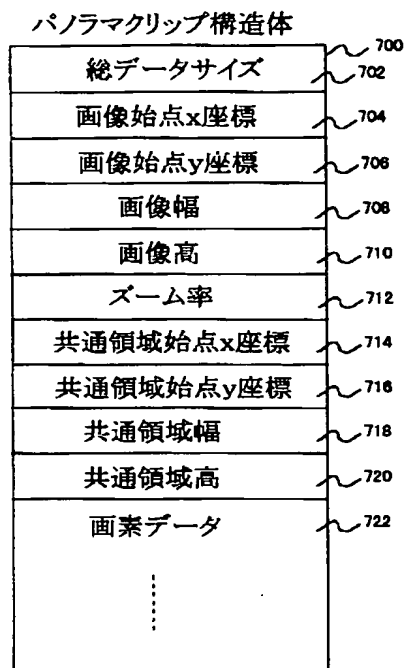


パノラマファイル構造体

ファイルID	600
パノラマ画像幅	602
パノラマ画像高	604
パノラマ画像色数	606
パノラマ画像フォーマット	608
パノラマ画像xオフセット	610
パノラマ画像yオフセット	612
パノラマクリップ構造体1	614
パノラマクリップ構造体2	700
...	700

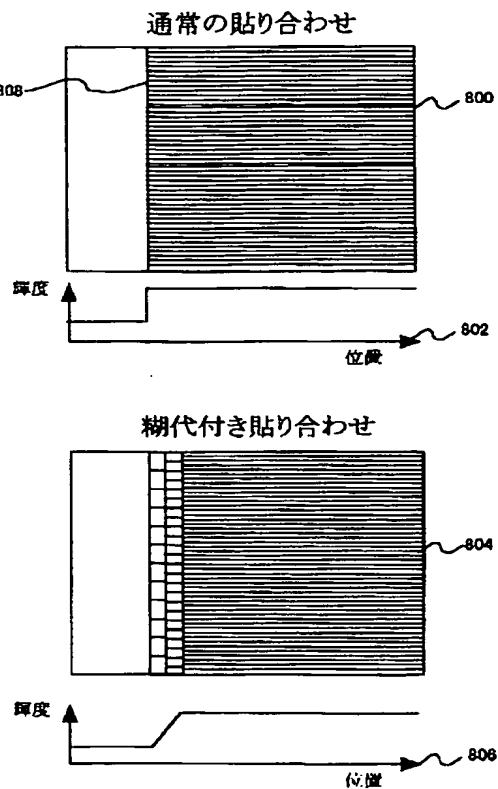
【図7】

図7



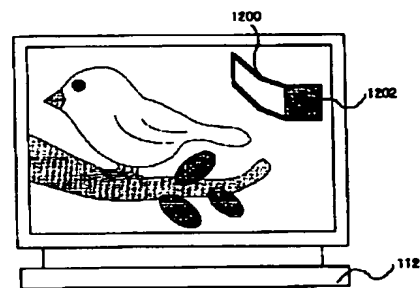
【図8】

図8



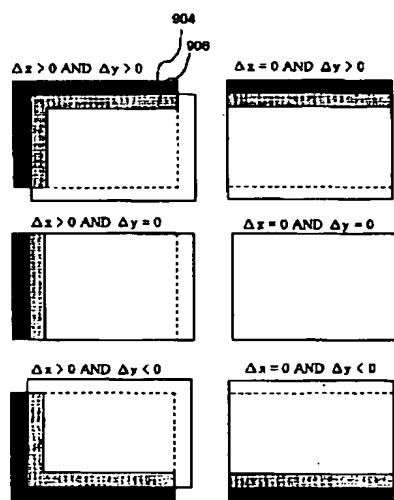
【図12】

図12



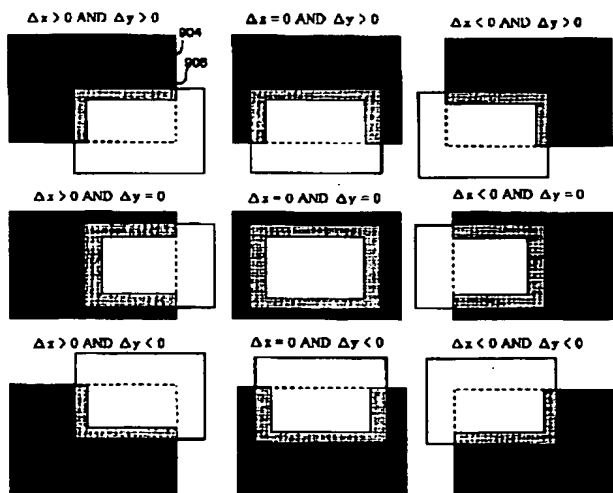
【図10】

図10



【図11】

図11



## フロントページの続き

Fターム(参考) 5B057 AA20 BA02 CD05 CE10 DC22  
DC25 DC30  
5C023 AA02 AA11 AA21 AA38 BA01  
BA11 BA15 CA01 CA03 CA08  
DA04 DA08 EA02

**\* NOTICES \***

**Japan Patent Office is not responsible for any damages caused by the use of this translation.**

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

**CLAIMS**

---

[Claim(s)]

[Claim 1] An image synthesizer unit characterized by providing the following A display means to display an animation under photography A means to capture a frame image one by one from said animation, to compare two frame images with which incorporation time of day differs, and to calculate movement magnitude A means to memorize a subtraction image and said movement magnitude of said two frame images A means to create indicator information which symbolizes a synthetic image compounded from memorized movement magnitude and a captured frame image, and a means to compound a panorama image using said subtraction image and said movement magnitude

[Claim 2] It is the image synthesizer unit which has a means compound a contraction image from a contraction frame image which said indicator information reduced said synthetic image, and reduced said captured frame image for a predetermined scale factor in an image synthesizer unit according to claim 1, and movement magnitude which changed said movement magnitude according to this scale factor, and is characterized by for said display means to display said contraction image and said panorama image.

[Claim 3] It is the image synthesizer unit which has a means to create a symbol image in which said indicator information is a symbol image which symbolizes said synthetic whole image configuration in an image synthesizer unit according to claim 1, and a configuration of said synthetic image is shown, and is characterized by said display means displaying said symbol image and said panorama image.

[Claim 4] An image synthesizer unit characterized by setting to an image synthesizer unit according to claim 2, and carrying out highlighting of the field corresponding to a frame incorporated by the newest in said contraction image.

[Claim 5] An image synthesizer unit characterized by setting to an image synthesizer

unit according to claim 3, and carrying out highlighting of the field corresponding to a frame incorporated by the newest in said symbol image.

[Claim 6] Said subtraction image is an image synthesizer unit characterized by said panorama image composition means compounding a boundary between said two frame images smoothly in an image synthesizer unit according to claim 1 using this common image field also including a common image field of said two frame images.

[Claim 7] It is the image synthesizer unit characterized by calculating a pixel of this expanded subtraction image with interpolation when expanding said subtraction image and creating a synthetic image in an image synthesizer unit according to claim 1.

[Claim 8] An image synthesizer unit characterized by providing the following. A display means to display a photoed animation A means to capture a frame image one by one from said animation A means to compare two frame images with which incorporated time of day differs, and to calculate movement magnitude of an image A means reduce this synthetic image remembered to be a means obtain a synthetic image combining a frame image captured based on said movement magnitude, a buffer means memorize said synthetic image or its contraction image, and a preview display means display one field of said synthetic image memorized by said buffer by said buffer when said synthetic image or its contraction image exceed a predetermined field, and said one field

[Claim 9] A record medium with which an image composition method characterized by providing the following was recorded and in which a computer readout is possible A step which displays an animation under photography A step which captures a frame image one by one from said animation, compares two frame images with which incorporation time of day differs, and calculates movement magnitude A step which memorizes a subtraction image and said movement magnitude of said two frame images A step which creates indicator information which symbolizes a synthetic image compounded from memorized movement magnitude and a captured frame image, and a step which compounds a panorama image using said subtraction image and said movement magnitude

[Claim 10] A step which compares two frame images with which incorporated time of day differs, and calculates movement magnitude of an image, A step which obtains a synthetic image combining a frame image captured based on this movement magnitude, A step which memorizes this synthetic image or its contraction image to a buffer, A step which displays one field of this synthetic image memorized by this buffer and which indicates by preview, A record medium with which an image composition method by which it is having [ when this synthetic image or its contraction image exceeds the field range of regular ]-step which reduced and displays this synthetic image and this one

field which were memorized by this buffer characterized was recorded and in which a computer readout is possible.

---

## DETAILED DESCRIPTION

---

### [Detailed Description of the Invention]

[0001]

[The technical field to which invention belongs] This invention relates to the image synthesizer unit which compounds a panorama image on real time from the dynamic image which is applied to the record medium with which the synthesizer unit and the synthetic method of compounding the panorama image of a wide-field-of-view angle were recorded, especially continues, and this record medium by which synthetic method drawing record was carried out.

[0002]

[Description of the Prior Art] Two or more images which photoed a different visual field are compounded, and the technology which makes the panorama image of the wide field of view of one sheet attracts attention. Especially recently, interests have gathered for the technology which creates a panorama image similarly from the continuation dynamic image photoed while carrying out the pan of the video camera and changing the visual field and which is called Video Mosaic. For example, Teodosio of U.S. Massachusetts Institute of Technology, L. Bender, W. It depends. Salient Video Stills: Content and Context Preserved, Proc. ACM Multimedia 93, pp. 39-46 (1993) etc. -- it is known. In addition, although various examples of research exist, the panorama image has been obtained by sticking an image so that the common area may be found and a common area may lap exactly about the image of two or more sheets with which any technique was acquired continuously. Since, as for close, the image of a new visual field comes to a photography image by following on moving a camera, they are connected and, finally the panorama image of the wide field of view is obtained.

[0003] However, in many examples of research, things for which the precision of a synthetic image is raised, such as how such a common area is found with a sufficient precision or how to amend a gap of the image periphery by distortion of a lens, are thought as important, and it is hard to say that the examination from a viewpoint of the usability of the operator using such panorama composition technology is actually enough. In order to perform image recognition generally about the natural image which cannot specify the class of target image, human being needs to intervene and it is necessary to give suitable decision on various aspects of affairs. In this case, the

high-speed responsibility of processing influences an operator's user-friendliness greatly. depending on a use, it connects that there is no breakdown as an image -- even having -- if it is, it is not few also when the technique excellent in usability which can process an input image without delay to the input of real time, i.e., an image, rather, and the result can check immediately is more desirable.

[0004]

[Problem(s) to be Solved by the Invention] Two kinds of processings need to be performed in creation of a panorama image. One is presumed processing of the amount of camera motions in which recognize a continuous input image and the coordinate of a common area is determined. Another is obtained synthetic processing which moves and actually connects an image from an amount. It is most desirable to perform these to coincidence on real time moreover from a viewpoint of usability. That is, when the shaken amount is calculated, an image ties, and is set automatically and it finishes photographing, it is an ideal that the synthetic image is already made and it can check, at the same time it shakes a camera. Redo can be done freely any number of times, even if it adopts how with missing information to essentially photograph an uncompoundable image, on that spot, a mistake is known and the fundamental correction from how to photograph can be performed, until the image included in mind is obtained by this.

[0005] When the shaken amount is calculated, an image ties, and is set automatically, and it finishes photographing, and the synthetic image is already made and it enables it to check at the same time it shakes a camera, both these two kinds of presumed processings and synthetic processing need to process on real time. About presumed processing of the amount of camera motions, there is high-speed mode of processing shown by JP,11-004398,A "a digital wide camera" by artificers, for example. However, about synthetic processing, it cannot be said that even the level which can be performed on real time at coincidence fully inquires. After photography, a synthetic result cannot be checked but it is inadequate that the amount of motions is just calculable on real time to photography and coincidence in the field of usability, if it does not wait for the time amount of lamination processing of images, although a calculated part of the amount of motions becomes short.

[0006] The purpose of this invention is to realize high-speed responsibility of the synthetic processing which can check a synthetic result on real time in order to raise an operator's usability in the method of compounding a panorama image from a continuous dynamic image.

[0007]

[Means for Solving the Problem] A display means to display an animation under



photography in order to attain the above-mentioned purpose, and a means to capture a frame image one by one from said animation, to compare two frame images with which incorporation time of day differs, and to calculate movement magnitude, A means to record a subtraction image and said movement magnitude of said two frame images, A means to create indicator information which symbolizes a synthetic image compounded from memorized movement magnitude and a captured frame image, and a means to compound a panorama image using said subtraction image and said movement magnitude are established.

[0008]

[Embodiment of the Invention] Hereafter, one example of this invention is explained to details. Drawing 1 is an example of the outline block diagram of the system configuration for realizing this invention. 100 is the general video camera image pick-up section, and transmits the photoed image to a processing unit 116 through the video signal line 106. The camera section is connected with a processing unit 116 by the control signal line 108, and 116 has come to be able to perform a setup of the camera parameter of being as adjusting a white balance \*\*\*\* [, and ], if needed. [ changing the zoom scale factor of a video camera ] It can omit, when such a setup is unnecessary. The processing unit 116 is the same as the system configuration of the digital computer used for the current general-purpose target, especially a personal computer, the display output is drawn on the screen of a display 112 through a video signal line, and carrier beam directions are notified to a processing unit 116 by the input device 120 from a user. Serially, the video signal outputted from the camera section 100 is changed into digital image data by the A/D converter of 102, goes into memory 114 through an interface 104, and is processed by CPU110 according to the program stored in another address of memory 114.

[0009] Here, although the signal from a video camera explains in the example in the case of analog video format, such as NTSC, when using the video camera of a digital output, the A/D converter of 102 is unnecessary. In this case, the converter of image format may be instead needed. Moreover, according to the necessity for processing, various information can be accumulated in the external information storage device of 118.

[0010] In case various kinds of data created by the processing explained below is stored in memory 114 and CPU110 processes, it is referred to if needed. Although the following explanation explains as software which CPU performs altogether, the part or all may be replaced and performed by the hardware logic circuit. Moreover, you may be the intelligent video camera gestalt which unified the camera section 100, the processing

unit 116, and the display 112.

[0011] Below, the software flow which is performed on the above-mentioned hardware and which solves the technical problem of point \*\* is explained to details. First, the creation method of the panorama image using the above-mentioned software is explained. A user tells shifting to the condition (it being called panoramic exposure mode below) that panorama composition is first performed on real time to a processing unit 116 with the input device 120. When an input device uses a computer system, in taking a camera one apparatus gestalt in a keyboard, a mouse, etc., switches, such as a "panoramic exposure initiation carbon button", correspond. If a processing unit 116 goes into panoramic exposure mode, it will carry out the \*\* frame input of the video signal outputted from a video camera, will calculate the amount of camera motions of the inputted frame image, and will collect parameters and data required for panorama composition. At this time, that whole or some fields are set up by the display 112 as a preview screen for checking a synthetic condition, and the appearance of the synthetic image connected according to the motion of a camera is indicated by updating on real time at it.

[0012] The image is shown in the display 112 of drawing 1 . In this invention, the synthetic image 122 (it is hereafter called a preview image) for a preview display is characterized by always carrying out a scaling and displaying that it becomes the greatest magnitude in which the whole fits in a preview screen. By this, the grasp at an instant of the whole synthetic condition is enabled. This is fill uped with the suitable background color, although the null portion 124 in which an image does not exist exists with the configuration of a synthetic image when a preview screen is a rectangle at this time. The location of the newest input frame in a synthetic image is surrounding, or changing a color and carrying out highlighting by the frame, and the role of the compass which does not puzzle the direction from which a camera should be moved is also given (126).

[0013] Again, with the input device 120, if a user directs termination in panoramic exposure mode, he stops a new frame image input and will be in a standby condition. The preview image updated by the last frame input by the display 112 is held, and a user looks at the image, judges whether the result of a request is obtained, if good, he creates the synthetic image (it is hereafter called a final output composition image) as a final output of the real size which is not an object for a preview, and saves at the external information recording device 118 etc.

[0014] An example of a flow chart which realizes the above-mentioned function to drawing 2 is shown. This algorithm makes the following three points the basic concept

in order to enable high-speed processing indispensable to the above-mentioned function.

[0015] (1) Create the synthetic image of contraction size only to previews apart from a final output composition image.

[0016] (2) Let the image memorized in the last composition images be only an indispensable field.

[0017] (3) Reconcile creation of the preview image of an arbitration configuration and arbitration size, and memory effectiveness.

[0018] (1) is a device for reducing the time amount concerning synthetic processing. It is not necessary to be the image of size with the created image not necessarily big generally, in order to judge whether it is a desired thing. So, at the time of photography, a final output composition image is not created but creates the synthetic image of contraction size separately as an object for a preview. It is made for a play view image to be mostly completed to coincidence with photography by this. Since composition is performed based on the camera motion estimate from which this preview image was obtained, near, distortion, a mistake, etc. are extremely reflected in an actual synthetic result as it is. If a user judges whether the desired image was obtained and is checked by this preview with right \*\*, he will create the final output composition image of actual magnitude anew. Since it can redo immediately and the time amount of composition of real size can be saved when it is judged as failure, an efficient activity can be performed.

[0019] (2) is a device for reducing the chart lasting time of the former data for a final output composition image. When (1) is adopted and a final output composition image is made to be created later, it is necessary to record temporarily the REMU image which is the former data. It becomes the inhibition factor of the high-speed processing by it being as that required memory space \*\*\*\*\*s \*\*\*\* [ , and ] to record the new frame image inputted one after another one after another. [ that it takes a data transfer time for a long time ]

[0020] Here, as shown in drawing 3 , a panorama composition image piles up every sheet of a continuous frame image, and is obtained. Drawing 3 is an example which creates a panorama image from a series of panimages, expresses typically the image which consists of a frame image with which the bottom continues, and shows the panorama image with which the bottom is compounded from it. Two or more lines which tie up the upper and lower sides express the correspondence relation of frame images. By the lower panorama image, the continuous field (it is hereafter called a common area) which is in agreement always in image by inter-frame exists, and redundancy is very high as data. Then, a panorama image is fully reproducible by combining only the field of the portion which excluded the this continuation inter-frame common area.

Therefore, if only the field which excluded this common area per frame is recorded, the amount of data which should record per frame is reduced sharply, a data transfer time will be shortened and its required memory space will also decrease. Although the whole image is needed only for superposition of the last frame, since it is only the 1 time uniquely, and it is the last and the following frame does not exist even if time amount cuts some in many to data transfer, the problem of not being of use for the processing for degree frame is not generated, either.

[0021] Although it is about (3) of the last, a preview image conforms to a motion of a camera out of photography further, it is not only obtained by photography termination and coincidence, but is updated synchronizing with a motion of real time, i.e., a camera, and a check is [ way ] easy for a user and it is user-friendly. However, the panorama image compounded becomes large in monotone as it moves a camera, and the configuration and magnitude become various depending on how to move a camera. In order to display a preview image, the panorama image compounded for a preview must always be created on the memory area, but in order to create the possible preview image of all configurations and magnitude on a memory area, a beforehand very large memory area must be taken.

[0022] Though a contraction frame image is used with the indicator of (1), the capacity of a required memory area is huge. And there will also be many memory portions which are not used in fact, and, as for an emergency, the utilization ratio of memory will become low. Of course, the direction from which a camera is moved first can be determined and only a part required for it can secure a memory area. However, by this method, even if the complicatedness which must show the direction from which a user moves a camera beforehand arises, furthermore mind changes a user on the way, it cannot respond at all. Moreover, finally it becomes the image of the magnitude of which, or a user cannot predict in many cases. To correspond to the synthetic image of unrestricted size and an unrestricted configuration is desired.

[0023] Here, the magnitude of the preview screen for displaying a preview image on a monitor is immobilization, and when the magnitude of a preview image becomes larger than the preview screen of the magnitude of immobilization, a contraction scaling is carried out and it is displayed to suit the screen size. Conversely, if it says, the number of pixels exceeding the resolution of a preview screen will become useless in case a preview image is displayed.

[0024] Then, the magnitude of a preview image is  $n$  times (length or ratio as horizontal length.) of the magnitude of a preview screen. When  $n \times n$  times are exceeded in area, the magnitude of a preview image is compulsorily reduced by  $1/n$  time (similarly in area

1/(mxm) twice). And when the magnitude of the reduced preview image exceeds n times of a preview screen again, it repeats as 1/m time contraction of a preview image is performed. The effect that the preview image which can restrict the memory area size for preview images to n or less times of the size of a preview screen, and is displayed on a preview screen by this does not have deterioration of resolution compared with the case where it has a memory area beyond it is acquired.

[0025] Based on the above thing, it returns to explanation of the flow chart of drawing 2. Various kinds of initialization is first performed by processing 200. For example, the memory area (it is hereafter called a preview buffer) of the magnitude which carried out magnitude of a preview screen in all directions [ n times as many as this / each ] is secured as a working buffer for preview image creation. A preview image is compounded in this preview buffer. the field called a preview viewing area is set to a preview buffer, and this field fills the magnitude of a preview screen exactly -- as -- always -- a scaling -- that is, zooming is carried out and it becomes the display of a preview screen.

[0026] A sequential preview viewing area is expanded so that the whole preview image may always be contained, as the field of the same magnitude as the preview screen of a buffer center section is set up as a preview viewing area at first and a preview image becomes large on a preview buffer. Subsequently, the initial value of the preview reduction percentage r which is a variable for preview image creation is calculated (202). r is the ratio of the magnitude of the frame image inputted from a camera, and the magnitude of a preview screen, and is a parameter adjusted so that an input frame image may be displayed to the limit of a preview screen in an initial state without a motion of a camera. For example, it is  $r = 1$  when magnitude of an input frame image and magnitude of a preview screen are made the same. It is set to 0.

[0027] In the continuing processing 204, the frame image of the one newest coma of the image photoed with the video camera is inputted. And movement magnitude detection processing of an image is performed, comparing between the frame images inputted one after another. Although this movement magnitude detection considers various methods, such as an optical flow, the method indicated by the "digital wide camera" of Japanese Patent Application No. 9-153303 is used for it as a method computable at a high speed here. This method enables high-speed processing very much, in order to calculate movement magnitude from projection distribution of an image.

[0028] In processing 206, the level projection distribution Px and the perpendicular projection distribution Py are created from this frame image. Level projection distribution is what took the average of the color and brightness about each line of the pixel which constitutes an image, and is the single dimension sequence of the average

only for a line count. Perpendicular projection distribution is what took the average of a color or brightness similarly about each train of an image, and is the single dimension sequence of the average only for the number of trains. For example, the list of the value of perpendicular projection distribution has the property shifted to the forward direction or hard flow corresponding to it, when an image moves to right and left. Perpendicular projection distribution is similarly shifted to order hard flow corresponding to vertical movement of an image. Therefore, the horizontal / perpendicular projection distribution created from a certain frame image are compared with the horizontal / perpendicular projection distribution created from the frame image inputted into the degree, and if it asks for which the sequence of the average shifted, respectively, the movement magnitude of level/perpendicular direction between the images of two sheets can be found. Although the list of the value of projection distribution is completely in agreement with neither a noise nor some deformation since it is image change, one of the two of two projection distribution to compare is shifted by delta pixel one by one, and is compared (all integers for which delta takes the range of  $-R$  to  $+R$ ).  $R$  obtains the value which shows the retrieval range set up beforehand, and delta when dissimilarity becomes low most as movement magnitude. Dissimilarity uses absolute total of the difference of the color of the pixel of a corresponding line or a train, or each average of brightness etc. (210-212). However, since the frame image in comparison with the past does not exist in the first image input, it returns to processing 204 (208).

[0029] In processing 214, from obtained level / the amount delta of vertical migration, it asks for the common area of the last frame image and the present frame image, and records only about the field of portions other than the common area in the last frame. By this, there is little storage capacity compared with the case where it records about all fields, and it ends.

[0030] Moreover, in processing 216, the movement magnitude for preview composition which multiplied the value of the obtained horizontal / the amount delta of vertical migration by  $r$  is calculated, and the image reduced by  $r$  times is drawn in piles in the location which shifted only  $\text{deltaxr}$  obtained above to the contraction frame image which drew on the preview buffer immediately before. Thus, the preview viewing area updated so that the expanded whole preview image might be included is displayed on a preview screen (218).

[0031] At this time, when the whole preview image is as Hasama to the limit of the edge of a preview buffer, (220) and the whole preview buffer are reduced by  $1/m$  time (222). At this time, a preview viewing area is similarly reduced by  $1/m$  time. And  $r$  multiplies by  $1/m$  similarly.

[0032] The typical flow of contraction processing of this preview buffer is shown in drawing 4. 400 shows the whole preview buffer and 402 is a preview viewing area. 404 shows a preview image. A location and magnitude change so that a preview viewing area may always be circumscribed to the whole preview image. A expresses the situation immediately after panoramic exposure initiation, and the inputted frame image is drawn by the center section. The preview image is extended, also follows a preview viewing area, and becomes large as a camera is moved (B). And since a buffer will be overflowed if a preview image reaches to the edge of a preview buffer as shown in C, the whole preview buffer is reduced by  $1/m$  time here (D). At this time, a preview viewing area is also reduced for the same scale factor. And if a preview image is extended further, drawing of an image will be continued according to it.

[0033] Here, if  $n=3$  and  $m=2$ , first, the center of a preview buffer and the center of a viewing area will be doubled, and the size of the preview viewing area when arriving at elongation and an edge after that according to a photography image will become always larger than the twice (a length ratio in every direction is twice) of the original size of a preview screen. therefore, this -- every the whole preview buffer  $1/2$  twice (it is a length ratio in every direction  $1/2$ ) -- even if -- the size of a preview viewing area is not less than the original size of a preview screen. In addition, when a preview viewing area reaches that of a buffer exactly at the time of  $n=3$  and  $m=2$ , a preview viewing area becomes twice the original size of a preview screen. Then, if a preview buffer is doubled  $1/2$ , the size of a preview viewing area will become the same as the original size of a preview screen.

[0034] The result which hardly changes to the case where secure a huge memory area, draw a preview image freely, and it thins out and displays on the size of a preview screen on it by this will be obtained only by preparing  $n$  times as much memory space as a preview screen. Moreover, since what is necessary is just to thin out at intervals of a pixel simply in the case of  $m=2$ , contraction can also be extremely performed at a high speed.

[0035] Here, it considers creating a panorama image 10 times the magnitude for example, of an input frame image. By the method of securing a memory area beforehand, without being based on the above-mentioned method, in order to preview the panorama image of this area, an about 400 times as many memory area as an input frame image is needed beforehand. Since this cannot predict in which direction a camera moves from the beginning, it is because it is necessary to secure the rectangle field which has one 10 times the length of this in each vertical and horizontal direction centering on the first input frame image, respectively. However, according to the method of this invention, it

is [ at one at most 9 times the magnitude of this ] good in the case of  $n=3$ . Since this invention guarantees unrestricted size by one 9 times this magnitude of this, the difference of a memory area with the conventional method is followed on the magnitude of the panorama image to compound increasing 20 times and 30 times, and becomes large exponentially.

[0036] Furthermore, in high resolving still picture composition processing in which the zoom image was used, the above-mentioned technique has an effect further. The principle of the high resolving still picture composition by the zoom is shown in drawing 5. This fundamental view is already mentioned reference Salient Video Stills. : It is shown in Content and Context Preserved. A big high resolution image can be obtained by making the frame image when zooming in most into criteria size by piling up other frame images according to the rate of a zoom. This high resolution image becomes highly minute little by little gradually as a periphery approaches a core, although a pixel only becomes large simply. The periphery is fading and its image photoed with the usual still camera is also close to the property of the high resolution image by which superposition composition was carried out, when a focus is doubled with a center. That is, an image equivalent to having copied using the image sensor of expensive high resolution will be obtained using the image sensor of a cheap low resolution.

[0037] However, by such conventional technique, the expansion processing time also becomes long and a real-time operation becomes difficult further rather than the time of panorama composition as superposition processing must be performed and a dilation ratio becomes large, after expanding the image at the time of zoom out. What is necessary is just to draw the first frame by actual size, and to pile up the rest, expanding a frame image, if a huge memory area can be prepared first in the case of an image which carries out zoom out in monotone still gradually. However, in the case of an image which zooms in reverse, it is necessary to expand the synthetic image, and time amount requires expanding a huge memory area the whole \*\* frame very much so that the frame image which is the purpose of high resolving image composition and into which resolution was inputted at the end by composition with slight height may always become actual size. This is the technical problem of the real-time operation no one but which cannot predict any motion [ what kind of camera ] there is beforehand.

[0038] On the other hand, if the view of this invention is applied, a preview can be checked on real time in few memory areas. About the technique of asking a high speed for the rate of a zoom, the technique already shown in Japanese Patent Application No. 10-53100 by artificers "a super resolution camera" can be used, for example. When it checks here, it is piling up the frame image inputted, adjusting size so that it may



become the contraction scale as composition of a high resolving image with the same common area between continuous frame images. Therefore, like the time of panorama composition, in order to obtain a final synthetic image, a common area requires only the character type outer frame portion [ in / in the case of the image of zoom-in / the last frame ] of RO except a common area that what is necessary is to record only once. Moreover, what is necessary is just to merely draw the frame image reduced according to the rate of a zoom one after another to the preview buffer in creation of a preview image.

[0039] The rate of a zoom is defined here as a ratio when carrying out size doubling so that a common area may be in agreement on the basis of the frame image inputted first. It is 1 when there is no zoom. In 0 and zoom-in, it is smaller than 1, and, in the case of zoom out, it becomes large from 1. On the other hand, in the case of the image of zoom out, since the range where the newest frame is larger is copied, the character type outer frame portion of RO except the common area in the newest frame is recorded shortly. If creation of a preview image hits, this outer frame portion is reduced according to the rate of a zoom, and it draws. If the rate of a zoom becomes [ the size of a preview image ] large, since the size of a preview buffer will be exceeded in the case of zoom out, the whole preview buffer is reduced to 1/m like the time of panorama composition in that case. Thus, an expansion required for preview image creation also in the case of a zoom is only processing to which only the character type outer frame portion of RO is expanded to n or less times, and can be clearly processed at a high speed. By the above, the result of high resolving image composition can also be checked on real time to photography and coincidence.

[0040] An example of the former data recorded in this example for a final synthetic image is shown in drawing 6 and drawing 7 . Data is divided roughly into the panorama file structure object 600 and the panorama clip structure 700. The former panorama file structure object is a header which gathers the whole former data, and the latter panorama clip structure is image data as which record is determined for every frame. The real size of a file ID 602 and a synthetic image is stored in a panorama file structure object in order of width of face (604) and height (606) as an identifier which shows that consecutiveness data is former data for image composition. Since the final size of a synthetic image can be expressed as the minimum rectangle circumscribed to it even if it was determined from the amount of camera motions and has taken the special configuration, only width of face and height are stored. Moreover, the color number 608 (it is one full color 24 bits etc. about whether they are for example, 256 colors) and the graphics format 610 (for example, it is the YUV format about whether it is RGB format)

of former data are also stored.

[0041] Moreover, the starting point coordinate of the rectangle field circumscribed to a synthetic image is stored as the panorama image x and y offset (612 614), and the coordinate of the panorama clip later mentioned on the basis of this coordinate is calculated. The panorama clip structure 700 is located in a line as data following this header information. The panorama clip structure serves as data for every one frame, and stores the location and size of the frame image in a synthetic image. First, 702 is the total data size and stores the amount of image data. The image starting point x-coordinate 702 expresses the x-coordinate value of the starting point of the frame image matched with this panorama clip structure when making into the starting point coordinate of a synthetic image the coordinate shown by the panorama image x above-mentioned offset 612. Similarly, the image starting point y-coordinate 704 expresses the y-coordinate value of the starting point of a frame image. The image width of face 706 and the image high 708 express real [ that a frame image is inputted ] size. The size of the frame image on a synthetic image is calculated by multiplying this by the rate 710 of a zoom. 714-720 specify the range of the above-mentioned common area in this frame image. The common area starting point x-coordinate 714 and the common area starting point y-coordinate 716 express the starting point coordinate of the common area in a frame, are the common area width of face 718 and common area high [ 720 ], and express the magnitude. All pixel data is located in a line with the pixel data 722 in order, scanning in the direction of the lower right from the upper left. When the suitable value for 714-720 is in close, the pixel data about a common area is disregarded simply.

[0042] Next, how to create a final synthetic image from the former data memorized above is explained. With reference to the panorama file structure object 600, the width of face and the height of a synthetic image are investigated first, and the rectangle image field of the magnitude is secured in a memory area. And with reference to the panorama clip structure 700 which follows, the location and magnitude which should draw are investigated, it asks for the relative position in the secured memory area, and image data 722 is written in. However, when at least one clip whose rate of a zoom is less than 1 time exists, proportionality conversion of the rate of a zoom of all the panorama clip structures is carried out so that the smallest rate of a zoom may be first set to 1, namely, so that the frame image which zoomed in most may be drawn by actual size. In this case, the size of a synthetic image also becomes large proportionally. It is lost by this that a frame image is reduced and composition is performed only in the direction in which resolution is raised.

[0043] Smoothing processing can be performed in order to show more beautifully the image composition described above. Generally, when sticking an image, there may be distortion in the time of the input image according to the aberration of a lens in that the amount of motions cannot be found correctly \*\*\*\*, or the brightness in a photography time may be changed, and a delicate boundary line may occur in the knot of an image. With human being's perception, this boundary line needs to amend in order to give an unpleasant impression. The easiest method of performing such amendment is sticking so that an edge's left for applying paste may be given among the images to stick, a boundary line's may be inserted and it may change to the color of the image of another side gradually from the color of one image. The image which was simplified very much and showed the concept is shown in drawing 8 . 800 shows signs that the image of two sheets is stuck simply. Here, in order to simplify explanation, the image of two sheets assumes that it is uniform by specific brightness, respectively. On the boundary line of an image, if change of the horizontal brightness at this time is looked at, as shown in a graph 802, it will change in an instant and this will form a boundary line 808. On the other hand, as shown in 804 and 806, if it sticks so that it may become a gradual change from the brightness of one image to the brightness of the image of another side, a boundary line will not be conspicuous [ an edge left for applying paste is given to the boundary of an image, and ].

[0044] In order to realize the above-mentioned smoothing processing, when recording former data by the synthetic method of this invention, only the part of an edge left for applying paste should record a larger field. By the increment in the necessary minimum amount of data, smoothing becomes possible by this. And the synthetic image connected smoothly is obtained by performing smoothing processing about the part of an edge left for applying paste in the case of next synthetic processing. The field made into an edge left for applying paste is chosen in the form which narrows the common area between the continuous images of two sheets fundamentally. The example in the case of having arranged so that a common area may lap with drawing 9 about the two frame images 900 and 902 is shown. Supposing 902 is a new frame image in time, the L character mold hatching field 904 at the upper left of drawing will be recorded as former image data. Since an edge left for applying paste should just bury the boundary line of two images gradually, it serves as a field which needs the field which swelled width of face for the direction which narrows a common area in accordance with a boundary line. In the example of drawing 9 , the field used as an edge left for applying paste turns into the hatching field 906 of the L character mold inside 904. It is shown in a list at drawing 10 by carrying out combination of the typical amount of camera motions, and the

edge-left-for-applying-paste field at that time.  $\Delta x$  and  $\Delta y$  show the movement magnitude of level/perpendicular direction among drawing, respectively. The pattern of the same edge-left-for-applying-paste field in case a zoom is further shown in drawing 11 is shown. By a diagram, the case of zoom-in is shown. Thus, even if there is a zoom, from the boundary line between the images when doubling magnitude so that the magnitude of a common area may become the same, it extends in the direction of a common area, and an edge-left-for-applying-paste field is decided. Also in zoom out, it is the same only by the relation to reverse of the frame images 900 and 902 becoming.

[0045] In addition, when creating a highly minute image, in case image data is stuck according to the rate of a zoom, it cannot be overemphasized that a still better-looking image is obtained by expanding using pixel interpolation technology, such as bilinear interpolation instead of expansion of a simple pixel. And a still more beautiful image is obtained by performing the above-mentioned smoothing processing for the image data expanded with pixel interpolation.

[0046] Moreover, although the technique of creating and displaying a preview image on real time was described above, as it indicates by simple instead of a preview image as the symbol which shows the situation under composition, coexistence with usability and much more improvement in the speed can also be aimed at. For example, since it is only a line drawing, drawing becomes possible at a high speed, at the same time necessity is also lost by reducing and sticking a frame image by showing with a line drawing whether not the preview image itself but the silhouette, i.e., a synthetic image, is made of what kind of configuration. The example of a preview screen based on this is shown in drawing 12. On a preview screen, an input image is displayed as it is, and an indicator indication also of the line drawing 1200 is given together with it on it. Moreover, in order to show where [ in a synthetic image ] the frame under current input hits, as shown in 1202, highlighting of the field which a frame image current in the inside of a line drawing occupies may be carried out. The purpose of checking a synthetic condition on real time is realized more by the operation engine performance of low cost by this.

[0047]

[Effect of the Invention] According to this invention, in creation of a panorama image or a high resolving image, the outstanding usability which can be retaken can be freely offered any number of times until the image which could check the advance condition of processing and the result to photography and coincidence, and went into them at mind is obtained.

---

## DESCRIPTION OF DRAWINGS

---

### [Brief Description of the Drawings]

[Drawing 1] It is an example of system configuration drawing for realizing this invention.

[Drawing 2] It is an example of the flow chart of the processing for realizing this invention.

[Drawing 3] It is image drawing which compounds a panorama image from the animation which carries out a pan.

[Drawing 4] It is an example of the contents of the preview buffer at the time of performing this invention.

[Drawing 5] It is image drawing which compounds a high resolving image from the animation which carries out a zoom.

[Drawing 6] It is an example of the data structure used by processing of this invention.

[Drawing 7] It is an example of the data structure used by processing of this invention.

[Drawing 8] It is image drawing showing the image lamination method.

[Drawing 9] It is drawing explaining each image field which processing of this invention defined.

[Drawing 10] It is an example of the pattern of an edge-left-for-applying-paste field used by processing of this invention.

[Drawing 11] It is an example of the pattern of an edge-left-for-applying-paste field used by processing of this invention.

[Drawing 12] It is an example of an indicator display used by processing of this invention.

### [Description of Notations]

100 [ -- A video signal line, 108 / -- A control signal line, 110 / -- CPU, 112 / -- A display, 114 / -- Memory, 116 / -- A processing unit, 118 / -- An external information storage device, 120 / -- Input device. ] -- The camera section, 102 -- An A/D converter, 104 -- An interface, 106